

SEQUENCE LISTING



<110> van Rooijen, Gijs  
Keon, Richard Glenn  
Boothe, Joseph  
Shen, Yin

<120> Commercial Production of Chymosin in Plants

<130> 9369-153

<140> 09/643,755

<141> 2000-08-23

<160> 4

<170> PatentIn Ver. 2.0

<210> 1

<211> 1173

<212> DNA

<213> Bovine

<220>

<221> CDS

<222> (1)..(1173)

<400> 1

atg aac ttc ctt aag tct ttc cct ttc tac gct ttc ctt tgt ttc ggt	48
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly	
1 5 10 15	
caa tac ttc gtt gct gtt act cac gct gct gag atc acc cgc att cct	96
Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro	
20 25 30	
ctc tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt	144
Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu	
35 40 45	
cta gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac	192
Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr	
50 55 60	
tcc ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat	240
Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp	
65 70 75 80	
agt caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc	288
Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe	
85 90 95	
acc gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc	336
Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile	
100 105 110	
tac tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga	384
Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg	
115 120 125	
aag tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac	432
Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr	
130 135 140	
ggt aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc	480

Gly 145	Thr	Gly	Ser	Met	Gln 150	Gly	Ile	Leu	Gly	Tyr 155	Asp	Thr	Val	Thr	Val 160	
tcc	aac	att	gtg	gac	att	caa	cag	aca	gta	gga	ctt	agc	acc	caa	gaa	528
Ser	Asn	Ile	Val	Asp 165	Ile	Gln	Gln	Thr	Val 170	Gly	Leu	Ser	Thr	Gln 175	Glu	
cca	ggt	gat	gtc	ttc	acc	tat	gca	gaa	ttc	gat	ggc	atc	ctt	ggt	atg	576
Pro	Gly	Asp	Val 180	Phe	Thr	Tyr	Ala	Glu 185	Phe	Asp	Gly	Ile	Leu 190	Gly	Met	
gca	tac	cca	tcg	ctc	gcg	tca	gag	tac	tcg	ata	cct	gtg	ttt	gac	aac	624
Ala	Tyr	Pro 195	Ser	Leu	Ala	Ser	Glu 200	Tyr	Ser	Ile	Pro	Val 205	Phe	Asp	Asn	
atg	atg	aac	cga	cac	cta	gta	gct	caa	gac	ttg	ttc	tcg	gtt	tac	atg	672
Met	Met	Asn 210	Arg	His	Leu	Val 215	Ala	Gln	Asp	Leu	Phe 220	Ser	Val	Tyr	Met	
gac	agg	aat	ggc	cag	gag	agc	atg	ctc	acg	ctt	gga	gct	att	gat	cca	720
Asp 225	Arg	Asn	Gly	Gln 230	Glu	Ser	Met	Leu	Thr	Leu 235	Gly	Ala	Ile	Asp	Pro 240	
tcc	tac	tac	aca	gga	tct	ctt	cac	tgg	gtt	cca	gtc	act	gtg	cag	cag	768
Ser	Tyr	Tyr	Thr 245	Gly	Ser	Leu	His	Trp	Val 250	Pro	Val	Thr	Val	Gln 255	Gln	
tac	tgg	caa	ttc	act	gtg	gac	agt	gtc	acc	atc	agc	ggt	gtg	gtt	gtt	816
Tyr	Trp	Gln 260	Phe	Thr	Val	Asp	Ser	Val 265	Thr	Ile	Ser	Gly	Val 270	Val	Val	
gca	tgt	gaa	ggt	gga	tgt	caa	gct	atc	ttg	gat	acc	ggt	acg	tcc	aag	864
Ala	Cys	Glu 275	Gly	Gly	Cys	Gln	Ala 280	Ile	Leu	Asp	Thr	Gly 285	Thr	Ser	Lys	
ctg	gtc	gga	cct	agc	agc	gac	att	ctc	aac	att	cag	caa	gct	att	gga	912
Leu 290	Val	Gly	Pro	Ser	Ser	Asp 295	Ile	Leu	Asn	Ile	Gln 300	Gln	Ala	Ile	Gly	
gcc	aca	cag	aac	cag	tac	ggt	gag	ttt	gac	ata	gat	tgc	gac	aac	ctt	960
Ala 305	Thr	Gln	Asn	Gln 310	Tyr	Gly	Glu	Phe	Asp 315	Ile	Asp	Cys	Asp	Asn	Leu 320	
agc	tac	atg	cct	aca	gtt	gtc	ttt	gag	atc	aac	ggc	aag	atg	tac	cca	1008
Ser	Tyr	Met	Pro	Thr 325	Val	Val	Phe	Glu	Ile 330	Asn	Gly	Lys	Met	Tyr 335	Pro	
ctg	acc	ccc	tcc	gcc	tat	acc	agc	cag	gat	caa	ggg	ttc	tgc	acc	agt	1056
Leu	Thr	Pro	Ser 340	Ala	Tyr	Thr	Ser	Gln 345	Asp	Gln	Gly	Phe	Cys 350	Thr	Ser	
gga	ttc	cag	agt	gag	aac	cat	tcc	cag	aaa	tgg	atc	ttg	gga	gat	gtg	1104
Gly	Phe	Gln 355	Ser	Glu	Asn	His	Ser 360	Gln	Lys	Trp	Ile	Leu 365	Gly	Asp	Val	
ttc	att	cgt	gag	tac	tac	agc	gtc	ttt	gac	agg	gcc	aac	aac	ctc	gtt	1152
Phe 370	Ile	Arg	Glu	Tyr	Tyr	Ser 375	Val	Phe	Asp	Arg	Ala 380	Asn	Asn	Leu	Val	
ggg	cta	gct	aaa	gca	atc	tga										1173
Gly 385	Leu	Ala	Lys	Ala	Ile 390											

<210> 2  
<211> 390



Met 1	Asn	Phe	Leu	Lys 5	Ser	Phe	Pro	Phe	Tyr 10	Ala	Phe	Leu	Cys	Phe 15	Gly
Gln	Tyr	Phe	Val 20	Ala	Val	Thr	His	Ala 25	Ala	Glu	Ile	Thr	Arg 30	Ile	Pro
Leu	Tyr	Lys 35	Gly	Lys	Ser	Leu	Arg 40	Lys	Ala	Leu	Lys	Glu 45	His	Gly	Leu
Leu	Glu 50	Asp	Phe	Leu	Gln	Lys 55	Gln	Gln	Tyr	Gly	Ile 60	Ser	Ser	Lys	Tyr
Ser 65	Gly	Phe	Gly	Glu	Val 70	Ala	Ser	Val	Pro	Leu 75	Thr	Asn	Tyr	Leu	Asp 80
Ser	Gln	Tyr	Phe	Gly 85	Lys	Ile	Tyr	Leu	Gly 90	Thr	Pro	Pro	Gln	Glu 95	Phe
Thr	Val	Leu	Phe 100	Asp	Thr	Gly	Ser	Ser 105	Asp	Phe	Trp	Val	Pro 110	Ser	Ile
Tyr	Cys	Lys 115	Ser	Asn	Ala	Cys	Lys 120	Asn	His	Gln	Arg	Phe 125	Asp	Pro	Arg
Lys	Ser 130	Ser	Thr	Phe	Gln	Asn 135	Leu	Gly	Lys	Pro	Leu 140	Ser	Ile	His	Tyr
Gly 145	Thr	Gly	Ser	Met	Gln 150	Gly	Ile	Leu	Gly	Tyr 155	Asp	Thr	Val	Thr	Val 160
Ser	Asn	Ile	Val	Asp 165	Ile	Gln	Gln	Thr	Val 170	Gly	Leu	Ser	Thr	Gln 175	Glu
Pro	Gly	Asp	Val 180	Phe	Thr	Tyr	Ala	Glu 185	Phe	Asp	Gly	Ile	Leu 190	Gly	Met
Ala	Tyr	Pro 195	Ser	Leu	Ala	Ser	Glu 200	Tyr	Ser	Ile	Pro	Val 205	Phe	Asp	Asn
Met	Met 210	Asn	Arg	His	Leu	Val 215	Ala	Gln	Asp	Leu	Phe 220	Ser	Val	Tyr	Met
Asp 225	Arg	Asn	Gly	Gln	Glu 230	Ser	Met	Leu	Thr	Leu 235	Gly	Ala	Ile	Asp	Pro 240
Ser	Tyr	Tyr	Thr	Gly 245	Ser	Leu	His	Trp	Val 250	Pro	Val	Thr	Val	Gln 255	Gln
Tyr	Trp	Gln	Phe 260	Thr	Val	Asp	Ser	Val 265	Thr	Ile	Ser	Gly	Val 270	Val	Val
Ala	Cys	Glu 275	Gly	Gly	Cys	Gln	Ala 280	Ile	Leu	Asp	Thr	Gly 285	Thr	Ser	Lys
Leu	Val 290	Gly	Pro	Ser	Ser	Asp 295	Ile	Leu	Asn	Ile	Gln 300	Gln	Ala	Ile	Gly
Ala 305	Thr	Gln	Asn	Gln	Tyr 310	Gly	Glu	Phe	Asp	Ile 315	Asp	Cys	Asp	Asn	Leu 320
Ser	Tyr	Met	Pro	Thr 325	Val	Val	Phe	Glu	Ile 330	Asn	Gly	Lys	Met	Tyr 335	Pro

Leu Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser  
 340 345 350  
 Gly Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val  
 355 360 365  
 Phe Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val  
 370 375 380  
 Gly Leu Ala Lys Ala Ile  
 385 390

<210> 3  
 <211> 3957  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <221> CDS  
 <222> (1554)..(2726)

<220>  
 <223> Description of Artificial Sequence: Phaseolin promoter- pre-pro-  
 chymosin-phaseolin terminator

<400> 3  
 ctgcaggaat tcattgtact ccaggtatca ttatagtga agttttggct ctctcgccgg 60  
 tggtttttta cctctattta aaggggtttt ccacctaaaa attctggtat cattctcact 120  
 ttacttgta ctttaatttc tcataatctt tggttgaaat tatcacgctt ccgcacacga 180  
 tatccctaca aatttattat ttgttaaaca ttttcaaacc gcataaaatt ttatgaagtc 240  
 ccgtctatct ttaatgtagt ctaacatttt catattgaaa tatataattt acttaatttt 300  
 agcgttggtg gaaagcataa agattttattc ttattcttct tcatataaat gtttaatat 360  
 caatataaac aaattcttta ccttaagaag gatttcccat tttatatttt aaaaatatat 420  
 ttatcaaata tttttcaacc acgtaaactc cataataata agttggttca aaagtaataa 480  
 aatttaactc cataattttt ttattcgact gatcttaaag caacaccag tgacacaact 540  
 agccattttt ttctttgaat aaaaaaatcc aattatcatt gtattttttt tatacaatga 600  
 aaatttcacc aaacaatcat ttgtggtatt tctgaagcaa gtcattgtat gcaaaattct 660  
 ataattccca ttgacacta cggaagtaac tgaagatctg cttttacatg cgagacacat 720  
 cttctaaagt aattttaata atagttacta tattcaagat ttcatatatc aaataactca 780  
 tattacttct aaaaaattaa ttagatataa ttaaaatatt acttttttaa ttttaagttt 840  
 aattgttgaa ttgtgacta ttgatttatt attctactat gtttaaattg ttttatagat 900  
 agtttaaagt aaatataagt aatgtagtag agtgtagag tgttacccta aaccataaac 960  
 tataacattt atggtggact aattttcata tatttcttat tgcttttacc ttttcttggt 1020  
 atgtaagtcc gtaactagaa ttacagtggg ttgccatggc actctgtggt cttttggttc 1080  
 atgcatgggt cttgcgcaag aaaaagacaa agaacaaaga aaaaagacaa aacagagaga 1140  
 caaaacgcaa tcacacaacc aactcaaatt agtcactggc tgatcaagat cgccgcgtcc 1200

atgtatgtct	aaatgccatg	caaagcaaca	cggtgcttaac	atgcacttta	aatggctcac	1260
ccatctcaac	ccacacacaa	acacattgcc	tttttcttca	tcataccac	aaccacctgt	1320
atatattcat	tctcttccgc	cacctcaatt	tcttcacttc	aacacacgtc	aacctgcata	1380
tgcgtgtcat	cccatgccca	aatctccatg	catgttccaa	ccaccttctc	tcttatataa	1440
tacctataaa	tacctcta	atcactcact	tctttcatca	tccatccatc	cagagtacta	1500
ctactctact	actataatac	cccaacccaa	ctcatattca	atactactct	act atg	1556
					Met	
					1	
aac ttc ctt aag tct ttc cct ttc tac gct ttc ctt tgt ttc ggt caa	1604					
Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly Gln						
5 10 15						
tac ttc gtt gct gtt act cac gct gct gag atc acc cgc att cct ctc	1652					
Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro Leu						
20 25 30						
tac aaa ggt aag tct ctc cgt aag gcg ctg aag gaa cat gga ctt cta	1700					
Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu Leu						
35 40 45						
gaa gac ttc ttg cag aaa caa cag tat ggc atc agc agc aag tac tcc	1748					
Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr Ser						
50 55 60 65						
ggc ttc ggt gaa gtt gct agc gtg cca ctt acc aac tac ctt gat agt	1796					
Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp Ser						
70 75 80						
caa tac ttt ggg aag atc tac ctc gga acc ccg cct caa gag ttc acc	1844					
Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe Thr						
85 90 95						
gtt ctc ttt gat act ggt tcc tct gac ttc tgg gtt ccc tct atc tac	1892					
Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile Tyr						
100 105 110						
tgc aag agc aat gcc tgc aag aac cac caa aga ttc gat ccg aga aag	1940					
Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg Lys						
115 120 125						
tcg tcc acc ttc cag aac tta ggc aaa ccc ttg tct ata cac tac ggt	1988					
Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr Gly						
130 135 140 145						
aca ggt agc atg caa gga atc tta ggc tat gat acc gtc act gtc tcc	2036					
Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val Ser						
150 155 160						
aac att gtg gac att caa cag aca gta gga ctt agc acc caa gaa cca	2084					
Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu Pro						
165 170 175						
ggg gat gtc ttc acc tat gca gaa ttc gat ggc atc ctt ggt atg gca	2132					
Gly Asp Val Phe Thr Tyr Ala Glu Phe Asp Gly Ile Leu Gly Met Ala						
180 185 190						
tac cca tcg ctc gcg tca gag tac tcg ata cct gtg ttt gac aac atg	2180					
Tyr Pro Ser Leu Ala Ser Glu Tyr Ser Ile Pro Val Phe Asp Asn Met						
195 200 205						

atg aac cga cac cta gta gct caa gac ttg ttc tcg gtt tac atg gac	2228
Met Asn Arg His Leu Val Ala Gln Asp Leu Phe Ser Val Tyr Met Asp	
210 215 220 225	
agg aat ggc cag gag agc atg ctc acg ctt gga gct att gat cca tcc	2276
Arg Asn Gly Gln Glu Ser Met Leu Thr Leu Gly Ala Ile Asp Pro Ser	
230 235 240	
tac tac aca gga tct ctt cac tgg gtt cca gtc act gtg cag cag tac	2324
Tyr Tyr Thr Gly Ser Leu His Trp Val Pro Val Thr Val Gln Gln Tyr	
245 250 255	
tgg caa ttc act gtg gac agt gtc acc atc agc ggt gtg gtt gtt gca	2372
Trp Gln Phe Thr Val Asp Ser Val Thr Ile Ser Gly Val Val Val Ala	
260 265 270	
tgt gaa ggt gga tgt caa gct atc ttg gat acc ggt acg tcc aag ctg	2420
Cys Glu Gly Gly Cys Gln Ala Ile Leu Asp Thr Gly Thr Ser Lys Leu	
275 280 285	
gtc gga cct agc agc gac att ctc aac att cag caa gct att gga gcc	2468
Val Gly Pro Ser Ser Asp Ile Leu Asn Ile Gln Gln Ala Ile Gly Ala	
290 295 300 305	
aca cag aac cag tac ggt gag ttt gac ata gat tgc gac aac ctt agc	2516
Thr Gln Asn Gln Tyr Gly Glu Phe Asp Ile Asp Cys Asp Asn Leu Ser	
310 315 320	
tac atg cct aca gtt gtc ttt gag atc aac ggc aag atg tac cca ctg	2564
Tyr Met Pro Thr Val Val Phe Glu Ile Asn Gly Lys Met Tyr Pro Leu	
325 330 335	
acc ccc tcc gcc tat acc agc cag gat caa ggg ttc tgc acc agt gga	2612
Thr Pro Ser Ala Tyr Thr Ser Gln Asp Gln Gly Phe Cys Thr Ser Gly	
340 345 350	
ttc cag agt gag aac cat tcc cag aaa tgg atc ttg gga gat gtg ttc	2660
Phe Gln Ser Glu Asn His Ser Gln Lys Trp Ile Leu Gly Asp Val Phe	
355 360 365	
att cgt gag tac tac agc gtc ttt gac agg gcc aac aac ctc gtt ggg	2708
Ile Arg Glu Tyr Tyr Ser Val Phe Asp Arg Ala Asn Asn Leu Val Gly	
370 375 380 385	
cta gct aaa gca atc tga agcttaataa gtatgaacta aaatgcatgt	2756
Leu Ala Lys Ala Ile	
390	
agggtgtaaga gctcatggag agcatggaat attgtatccg accatgtaac agtataataa	2816
ctgagctcca tctcacttct tctatgaata aacaaaggat gttatgatat attaacactc	2876
tatctatgca ccttattgtt ctatgataaa tttcctctta ttattataaa tcatctgaat	2936
cgtgacggct tatggaatgc ttcaaatagt acaaaaaacaa atgtgtacta taagactttc	2996
taaacaattc taacttttagc attgtgaacg agacataagt gttaagaaga cataacaatt	3056
ataatggaag aagtttgtct ccattttatat attatatatt acccacttat gtattatatt	3116
aggatgttaa ggagacataa caattataaa gagagaagtt tgtatccatt tatatattat	3176
atactaccca tttatatatt atacttatcc acttatttaa tgtctttata aggtttgatc	3236
catgatattt ctaatatattt agttgatatg tatatgaaag ggtactattt gaactctctt	3296

actctgtata aaggttggat catccttaaa gtgggtctat ttaattttat tgctttcttac 3356  
agataaaaaa aaaattatga gttgggttga taaaatattg aaggatttaa aataataata 3416  
aataataaat aacatataat atatgtatat aaatttatta taatataaca tttatctata 3476  
aaaaagtaaa tattgtcata aatctataca atcgtttagc cttgctggac gactctcaat 3536  
tatttaaacy agagtaaaca tatttgactt tttgggtatt taacaaatta ttatttaaca 3596  
ctatatgaaa tttttttttt ttatcggcaa ggaaataaaa ttaaattagg agggacaatg 3656  
gtgtgtccca atccttatac aaccaacttc cacaggaagg tcaggtcggg gacaacaaaa 3716  
aaacaggcaa gggaaatttt ttaatttggg ttgtcttggt tgctgcataa tttatgcagt 3776  
aaaacactac acataaccct tttagcagta gagcaatggt tgaccgtgtg cttagcttct 3836  
tttattttat ttttttatca gcaaagaata aataaaataa aatgagacac ttcagggatg 3896  
tttcaaccct tatacaaaaac cccaaaaaca agtttcttag caccctacca actaaggtag 3956  
c 3957

<210> 4  
<211> 390  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Phaseolin promoter- pre-pro-  
chymosin-phaseolin terminator

<400> 4  
Met Asn Phe Leu Lys Ser Phe Pro Phe Tyr Ala Phe Leu Cys Phe Gly  
1 5 10 15  
Gln Tyr Phe Val Ala Val Thr His Ala Ala Glu Ile Thr Arg Ile Pro  
20 25 30  
Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly Leu  
35 40 45  
Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys Tyr  
50 55 60  
Ser Gly Phe Gly Glu Val Ala Ser Val Pro Leu Thr Asn Tyr Leu Asp  
65 70 75 80  
Ser Gln Tyr Phe Gly Lys Ile Tyr Leu Gly Thr Pro Pro Gln Glu Phe  
85 90 95  
Thr Val Leu Phe Asp Thr Gly Ser Ser Asp Phe Trp Val Pro Ser Ile  
100 105 110  
Tyr Cys Lys Ser Asn Ala Cys Lys Asn His Gln Arg Phe Asp Pro Arg  
115 120 125  
Lys Ser Ser Thr Phe Gln Asn Leu Gly Lys Pro Leu Ser Ile His Tyr  
130 135 140  
Gly Thr Gly Ser Met Gln Gly Ile Leu Gly Tyr Asp Thr Val Thr Val  
145 150 155 160  
Ser Asn Ile Val Asp Ile Gln Gln Thr Val Gly Leu Ser Thr Gln Glu

